

1
1AP20 Rec'd PCT/PTO 13 APR 2006

SEQUENCE LISTING

<110> CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

<120> METHOD FOR THE DISSOCIATION OF THE EXTRACELLULAR HAEMOGLOBIN
MOLECULE OF ARENICOLA MARINA, CHARACTERIZATION OF THE PROTEIN
CHAINS CONSTITUTING SAID MOLECULE AND THE NUCLEOTIDE SEQUENCES
ENCODING SAID PROTEIN CHAINS

<130> WOB CNR GLOB

<150> FR 03/11992

<151> 2003-10-14

<160> 31

<170> PatentIn version 3.1

<210> 1

<211> 474

<212> DNA

<213> Arenicola marina

<220>

<221> CDS

<222> (1)..(474)

<223>

<400> 1

atg	aag	tcc	ttg	gtg	gtt	ctg	ttc	gcc	ctg	gtg	gcc	atg	gtg	gct	gca	48
Met	Lys	Ser	Leu	Val	Val	Leu	Phe	Ala	Leu	Val	Ala	Met	Val	Ala	Ala	
1				5					10					15		

gag	tgc	ggc	ccc	atg	cag	cgc	ctc	ctg	gtc	aag	acc	cag	tgg	aac	aag	96
Glu	Cys	Gly	Pro	Met	Gln	Arg	Leu	Leu	Val	Lys	Thr	Gln	Trp	Asn	Lys	
			20				25						30			

gtg	tac	ggc	acc	agc	aag	gtc	agg	gac	gag	gcc	gga	cac	gtc	ctc	tgg	144
Val	Tyr	Gly	Thr	Ser	Lys	Val	Arg	Asp	Glu	Ala	Gly	His	Val	Leu	Trp	
		35					40					45				

aag	gct	att	ttc	gcc	cag	gat	ccc	gag	acc	cgg	gct	ctc	ttc	aag	aga	192
Lys	Ala	Ile	Phe	Ala	Gln	Asp	Pro	Glu	Thr	Arg	Ala	Leu	Phe	Lys	Arg	
	50					55					60					

gtc	aac	ggg	gac	gac	atc	tac	tct	ccc	gag	ttc	atg	gct	cac	agc	gcc	240
Val	Asn	Gly	Asp	Asp	Ile	Tyr	Ser	Pro	Glu	Phe	Met	Ala	His	Ser	Ala	
65					70				75					80		

cgt	gtc	ttg	ggg	ggc	ctt	gac	att	gcc	atc	tcc	ctc	ctc	gac	aac	cag	288
Arg	Val	Leu	Gly	Gly	Leu	Asp	Ile	Ala	Ile	Ser	Leu	Leu	Asp	Asn	Gln	
			85					90					95			

gct	gac	ctt	gac	gtc	gcc	ctg	gct	cac	ctt	cac	gtg	cag	cac	gta	gaa	336
Ala	Asp	Leu	Asp	Val	Ala	Leu	Ala	His	Leu	His	Val	Gln	His	Val	Glu	
		100					105					110				

agg	cac	atc	cca	acc	cgc	tac	ttc	gat	ctg	ttc	aag	aac	gcc	ctg	atg	384
Arg	His	Ile	Pro	Thr	Arg	Tyr	Phe	Asp	Leu	Phe	Lys	Asn	Ala	Leu	Met	
		115					120					125				

gag tat gcc ccc agc gcc ctg gga cgc tgc ttc gat aag acc gcc tgg 432
 Glu Tyr Ala Pro Ser Ala Leu Gly Arg Cys Phe Asp Lys Thr Ala Trp
 130 135 140

agc tcg tgc ttt gac gtc atc gcc aac ggc atc aag gaa tag 474
 Ser Ser Cys Phe Asp Val Ile Ala Asn Gly Ile Lys Glu
 145 150 155

<210> 2
 <211> 157
 <212> PRT
 <213> Arenicola marina

<400> 2
 Met Lys Ser Leu Val Val Leu Phe Ala Leu Val Ala Met Val Ala Ala
 1 5 10 15
 Glu Cys Gly Pro Met Gln Arg Leu Leu Val Lys Thr Gln Trp Asn Lys
 20 25 30
 Val Tyr Gly Thr Ser Lys Val Arg Asp Glu Ala Gly His Val Leu Trp
 35 40 45
 Lys Ala Ile Phe Ala Gln Asp Pro Glu Thr Arg Ala Leu Phe Lys Arg
 50 55 60
 Val Asn Gly Asp Asp Ile Tyr Ser Pro Glu Phe Met Ala His Ser Ala
 65 70 75 80
 Arg Val Leu Gly Gly Leu Asp Ile Ala Ile Ser Leu Leu Asp Asn Gln
 85 90 95
 Ala Asp Leu Asp Val Ala Leu Ala His Leu His Val Gln His Val Glu
 100 105 110
 Arg His Ile Pro Thr Arg Tyr Phe Asp Leu Phe Lys Asn Ala Leu Met
 115 120 125
 Glu Tyr Ala Pro Ser Ala Leu Gly Arg Cys Phe Asp Lys Thr Ala Trp
 130 135 140
 Ser Ser Cys Phe Asp Val Ile Ala Asn Gly Ile Lys Glu
 145 150 155

<210> 3
 <211> 477
 <212> DNA
 <213> Arenicola marina

<220>
 <221> CDS
 <222> (1)..(477)
 <223>

<400> 3
 atg aag ttc ttg gtg gtt ctg ttc gcc ctg gtg gcc atg gtc gct gct 48
 Met Lys Phe Leu Val Val Leu Phe Ala Leu Val Ala Met Val Ala Ala
 1 5 10 15

gat tgt ggc ccc atg cag cgc ctc ctg gtc aag gcc cag tgg aac aag 96
 Asp Cys Gly Pro Met Gln Arg Leu Leu Val Lys Ala Gln Trp Asn Lys
 20 25 30

gtg tac ggc acc agc aag gtc agg gac gac gcc gga cac gtc ctc tgg 144
 Val Tyr Gly Thr Ser Lys Val Arg Asp Asp Ala Gly His Val Leu Trp
 35 40 45

aag gct atc ttc aac cag gat ggt gag acc cgc gcc ctc ttc aac aga 192
 Lys Ala Ile Phe Asn Gln Asp Gly Glu Thr Arg Ala Leu Phe Asn Arg
 50 55 60

gtg cac ggt gac gac atc tac tct ccc gag ttc atg gct cac agc gcc 240
 Val His Gly Asp Asp Ile Tyr Ser Pro Glu Phe Met Ala His Ser Ala
 65 70 75 80

cgt gtc ttg ggt ggc ctt gac att gcc atc tcc ctc ctc gac aac cag 288
 Arg Val Leu Gly Gly Leu Asp Ile Ala Ile Ser Leu Leu Asp Asn Gln
 85 90 95

gct gag ctt gac gct gtc ctg gct cac ctc aag gag cag cac att gag 336
 Ala Glu Leu Asp Ala Val Leu Ala His Leu Lys Glu Gln His Ile Glu
 100 105 110

agg ggg atc cca gac cgt tac ttc gac ctg ttc aag aac gcc ctg atg 384
 Arg Gly Ile Pro Asp Arg Tyr Phe Asp Leu Phe Lys Asn Ala Leu Met
 115 120 125

gag ttt gcc ccc agc gcc ttg gga cgc tgc ttc ata aag gac gct tgg 432
 Glu Phe Ala Pro Ser Ala Leu Gly Arg Cys Phe Ile Lys Asp Ala Trp
 130 135 140

agc tca tgc ttt gac gtc att gcc aac ggc atc aag gga cag taa 477
 Ser Ser Cys Phe Asp Val Ile Ala Asn Gly Ile Lys Gly Gln
 145 150 155

<210> 4

<211> 158

<212> PRT

<213> Arenicola marina

<400> 4

Met Lys Phe Leu Val Val Leu Phe Ala Leu Val Ala Met Val Ala Ala
 1 5 10 15

Asp Cys Gly Pro Met Gln Arg Leu Leu Val Lys Ala Gln Trp Asn Lys
 20 25 30

Val Tyr Gly Thr Ser Lys Val Arg Asp Asp Ala Gly His Val Leu Trp
 35 40 45

Lys Ala Ile Phe Asn Gln Asp Gly Glu Thr Arg Ala Leu Phe Asn Arg
 50 55 60

Val His Gly Asp Asp Ile Tyr Ser Pro Glu Phe Met Ala His Ser Ala
 65 70 75 80

Arg Val Leu Gly Gly Leu Asp Ile Ala Ile Ser Leu Leu Asp Asn Gln
 85 90 95

Ala Glu Leu Asp Ala Val Leu Ala His Leu Lys Glu Gln His Ile Glu
 100 105 110
 Arg Gly Ile Pro Asp Arg Tyr Phe Asp Leu Phe Lys Asn Ala Leu Met
 115 120 125
 Glu Phe Ala Pro Ser Ala Leu Gly Arg Cys Phe Ile Lys Asp Ala Trp
 130 135 140
 Ser Ser Cys Phe Asp Val Ile Ala Asn Gly Ile Lys Gly Gln
 145 150 155

<210> 5
 <211> 474
 <212> DNA
 <213> Arenicola marina

<220>
 <221> CDS
 <222> (1)..(474)
 <223>

<400> 5
 atg aag gtc ctg atc gta ctg atg gcc tgc ttg gcc tac gtc gcc gcc 48
 Met Lys Val Leu Ile Val Leu Met Ala Cys Leu Ala Tyr Val Ala Ala
 1 5 10 15
 gac tgc gga cct ctg cag agg ctg aag gtg aag cat cag tgg gtg cag 96
 Asp Cys Gly Pro Leu Gln Arg Leu Lys Val Lys His Gln Trp Val Gln
 20 25 30
 gtg tac agc ggc cat ggt tac gag cgt gag gcg ttc ggc aga gag gtc 144
 Val Tyr Ser Gly His Gly Tyr Glu Arg Glu Ala Phe Gly Arg Glu Val
 35 40 45
 ttc ctc gag atg tac aac cag gca ccc aag gcc aag gac ctc ttc acc 192
 Phe Leu Glu Met Tyr Asn Gln Ala Pro Lys Ala Lys Asp Leu Phe Thr
 50 55 60
 agg gtc agg ggc gag aac gtc ttc tcc ccc gag ttc gga gcc cac atg 240
 Arg Val Arg Gly Glu Asn Val Phe Ser Pro Glu Phe Gly Ala His Met
 65 70 75 80
 gtc cgt gtg ctc gga gga ctc gac atg tgc atc gct ctg ctg tcc gat 288
 Val Arg Val Leu Gly Gly Leu Asp Met Cys Ile Ala Leu Leu Ser Asp
 85 90 95
 gac acc gtc ctc aac gcc cag ctt gct cac ctc agc acg cag cac aag 336
 Asp Thr Val Leu Asn Ala Gln Leu Ala His Leu Ser Thr Gln His Lys
 100 105 110
 gac cgt gga atc ccc aac gag tac ttc gat gtg atg aag gtc gcc ctc 384
 Asp Arg Gly Ile Pro Asn Glu Tyr Phe Asp Val Met Lys Val Ala Leu
 115 120 125
 atg aag gtc gtc ccc ggc cac gtt tca cac ttc gac ttc gat gcc tgg 432
 Met Lys Val Val Pro Gly His Val Ser His Phe Asp Phe Asp Ala Trp
 130 135 140

tct gcc tgc tat gac gtc atc gcc aac ggc atc aag cac taa
 Ser Ala Cys Tyr Asp Val Ile Ala Asn Gly Ile Lys His
 145 150 155

474

<210> 6
 <211> 157
 <212> PRT
 <213> Arenicola marina

<400> 6
 Met Lys Val Leu Ile Val Leu Met Ala Cys Leu Ala Tyr Val Ala Ala
 1 5 10 15
 Asp Cys Gly Pro Leu Gln Arg Leu Lys Val Lys His Gln Trp Val Gln
 20 25 30
 Val Tyr Ser Gly His Gly Tyr Glu Arg Glu Ala Phe Gly Arg Glu Val
 35 40 45
 Phe Leu Glu Met Tyr Asn Gln Ala Pro Lys Ala Lys Asp Leu Phe Thr
 50 55 60
 Arg Val Arg Gly Glu Asn Val Phe Ser Pro Glu Phe Gly Ala His Met
 65 70 75 80
 Val Arg Val Leu Gly Gly Leu Asp Met Cys Ile Ala Leu Leu Ser Asp
 85 90 95
 Asp Thr Val Leu Asn Ala Gln Leu Ala His Leu Ser Thr Gln His Lys
 100 105 110
 Asp Arg Gly Ile Pro Asn Glu Tyr Phe Asp Val Met Lys Val Ala Leu
 115 120 125
 Met Lys Val Val Pro Gly His Val Ser His Phe Asp Phe Asp Ala Trp
 130 135 140
 Ser Ala Cys Tyr Asp Val Ile Ala Asn Gly Ile Lys His
 145 150 155

<210> 7
 <211> 498
 <212> DNA
 <213> Arenicola marina

<220>
 <221> CDS
 <222> (1)..(498)
 <223>

<400> 7
 atg ctt cgt ttc gta gca ctc ttg gct ctg gtc ggc ctg gcc gtc tgt
 Met Leu Arg Phe Val Ala Leu Leu Ala Leu Val Gly Leu Ala Val Cys
 1 5 10 15
 gac gac tgt tgt acc acc gag gac cgc aag gag gtc cag acg ctg tgg
 Asp Asp Cys Cys Thr Thr Glu Asp Arg Lys Glu Val Gln Thr Leu Trp
 20 25 30

48

96

agt gag atc tgg agt gcc cag ttc act ggt cgc cgt gtc cag gtt gcc 144
 Ser Glu Ile Trp Ser Ala Gln Phe Thr Gly Arg Arg Val Gln Val Ala
 35 40 45
 cag gct gtg ttc gag gac ctc ttc cgc cgc gac ccc gag tcc aag aac 192
 Gln Ala Val Phe Glu Asp Leu Phe Arg Arg Asp Pro Glu Ser Lys Asn
 50 55 60
 ctg ttc aag cgc gtc aat gtt gac gac atg aac agc ccc gaa ttc cac 240
 Leu Phe Lys Arg Val Asn Val Asp Asp Met Asn Ser Pro Glu Phe His
 65 70 75 80
 gct cac tgc atc cgt gtt gtc aac ggt ctt gac acc gtg atc ggt ctc 288
 Ala His Cys Ile Arg Val Val Asn Gly Leu Asp Thr Val Ile Gly Leu
 85 90 95
 ctt gac gac ccc gac acc ctg aag tcc cag ctc gag cac ttg gcc cag 336
 Leu Asp Asp Pro Asp Thr Leu Lys Ser Gln Leu Glu His Leu Ala Gln
 100 105 110
 cag cac aag gag cgt gat ggc atc cac aag acc cac ttc gac gag atg 384
 Gln His Lys Glu Arg Asp Gly Ile His Lys Thr His Phe Asp Glu Met
 115 120 125
 tcc cac gcc ttc ggc gcc gtc atg ccc cag gtc agc agc tgc ttc aac 432
 Ser His Ala Phe Gly Ala Val Met Pro Gln Val Ser Ser Cys Phe Asn
 130 135 140
 ccc gat gcc tgg aac cgt tgc ttc ggc tcc atc gct acc aag att gct 480
 Pro Asp Ala Trp Asn Arg Cys Phe Gly Ser Ile Ala Thr Lys Ile Ala
 145 150 155 160
 tcc ctc ctc gag gat taa 498
 Ser Leu Leu Glu Asp
 165

<210> 8
 <211> 165
 <212> PRT
 <213> Arenicola marina

<400> 8
 Met Leu Arg Phe Val Ala Leu Leu Ala Leu Val Gly Leu Ala Val Cys
 1 5 10 15
 Asp Asp Cys Cys Thr Thr Glu Asp Arg Lys Glu Val Gln Thr Leu Trp
 20 25 30
 Ser Glu Ile Trp Ser Ala Gln Phe Thr Gly Arg Arg Val Gln Val Ala
 35 40 45
 Gln Ala Val Phe Glu Asp Leu Phe Arg Arg Asp Pro Glu Ser Lys Asn
 50 55 60
 Leu Phe Lys Arg Val Asn Val Asp Asp Met Asn Ser Pro Glu Phe His
 65 70 75 80
 Ala His Cys Ile Arg Val Val Asn Gly Leu Asp Thr Val Ile Gly Leu
 85 90 95

Leu Asp Asp Pro Asp Thr Leu Lys Ser Gln Leu Glu His Leu Ala Gln
 100 105 110
 Gln His Lys Glu Arg Asp Gly Ile His Lys Thr His Phe Asp Glu Met
 115 120 125
 Ser His Ala Phe Gly Ala Val Met Pro Gln Val Ser Ser Cys Phe Asn
 130 135 140
 Pro Asp Ala Trp Asn Arg Cys Phe Gly Ser Ile Ala Thr Lys Ile Ala
 145 150 155 160
 Ser Leu Leu Glu Asp
 165

<210> 9
 <211> 498
 <212> DNA
 <213> *Arenicola marina*

<220>
 <221> CDS
 <222> (1)..(498)
 <223>

<400> 9
 atg atg tcc gtc gtg ttc ctc ctc ggc ctt gtg gcc tac gcc tcc gcc 48
 Met Met Ser Val Val Phe Leu Leu Gly Leu Val Ala Tyr Ala Ser Ala
 1 5 10 15
 tcc agc tgc tgc tcc tat gga gac cag cag aag gtc aag gcc cag tgg 96
 Ser Ser Cys Cys Ser Tyr Gly Asp Gln Gln Lys Val Lys Ala Gln Trp
 20 25 30
 aac agc ctc tgg aac acc cct gac tcc tcc aca tcc aag atc atc ttc 144
 Asn Ser Leu Trp Asn Thr Pro Asp Ser Ser Thr Ser Lys Ile Ile Phe
 35 40 45
 gga aag gaa gtc ttc gca cgc ttc ttc gag gtt gac ccc gag agc aag 192
 Gly Lys Glu Val Phe Ala Arg Phe Phe Glu Val Asp Pro Glu Ser Lys
 50 55 60
 agc ctg ttc ggt cgc gtc aag gtt gaa gac ccc gac agc ccc gag ttc 240
 Ser Leu Phe Gly Arg Val Lys Val Glu Asp Pro Asp Ser Pro Glu Phe
 65 70 75 80
 gcc gga cac gtg atc cgt gtt ttg acc ggt ctg gat ttg atc atc aac 288
 Ala Gly His Val Ile Arg Val Leu Thr Gly Leu Asp Leu Ile Ile Asn
 85 90 95
 ttg atg ggt gac gat gcc atg gat gcc gag ctg gcc cac ctt aac acc 336
 Leu Met Gly Asp Asp Ala Met Asp Ala Glu Leu Ala His Leu Asn Thr
 100 105 110
 cag cat ttg gcc aga gag gga atc acc gga acc cac ttc acc gag atg 384
 Gln His Leu Ala Arg Glu Gly Ile Thr Gly Thr His Phe Thr Glu Met
 115 120 125

ttc aag gtc ctg gat gga tcc ctc cgc cag gtt ctc gag gag tac gat 432
 Phe Lys Val Leu Asp Gly Ser Leu Arg Gln Val Leu Glu Glu Tyr Asp
 130 135 140

tcc ctg tcc tgg agg tac tgc ttc cgt ggt ctg ggc gcc gcc ctc agg 480
 Ser Leu Ser Trp Arg Tyr Cys Phe Arg Gly Leu Gly Ala Ala Leu Arg
 145 150 155 160

gat ggt ctc ccc gca taa 498
 Asp Gly Leu Pro Ala
 165

<210> 10
 <211> 165
 <212> PRT
 <213> Arenicola marina

<400> 10
 Met Met Ser Val Val Phe Leu Leu Gly Leu Val Ala Tyr Ala Ser Ala
 1 5 10 15

Ser Ser Cys Cys Ser Tyr Gly Asp Gln Gln Lys Val Lys Ala Gln Trp
 20 25 30

Asn Ser Leu Trp Asn Thr Pro Asp Ser Ser Thr Ser Lys Ile Ile Phe
 35 40 45

Gly Lys Glu Val Phe Ala Arg Phe Phe Glu Val Asp Pro Glu Ser Lys
 50 55 60

Ser Leu Phe Gly Arg Val Lys Val Glu Asp Pro Asp Ser Pro Glu Phe
 65 70 75 80

Ala Gly His Val Ile Arg Val Leu Thr Gly Leu Asp Leu Ile Ile Asn
 85 90 95

Leu Met Gly Asp Asp Ala Met Asp Ala Glu Leu Ala His Leu Asn Thr
 100 105 110

Gln His Leu Ala Arg Glu Gly Ile Thr Gly Thr His Phe Thr Glu Met
 115 120 125

Phe Lys Val Leu Asp Gly Ser Leu Arg Gln Val Leu Glu Glu Tyr Asp
 130 135 140

Ser Leu Ser Trp Arg Tyr Cys Phe Arg Gly Leu Gly Ala Ala Leu Arg
 145 150 155 160

Asp Gly Leu Pro Ala
 165

<210> 11
 <211> 771
 <212> DNA
 <213> Arenicola marina

<220>
 <221> CDS

<222> (1)..(771)

<223>

<400> 11

atg aag agc tac gtg ctc gtg tgc tgc ctc gtg gtg ggg gcc gtg gcc	48
Met Lys Ser Tyr Val Leu Val Cys Cys Leu Val Val Gly Ala Val Ala	
1 5 10 15	
tac ccc cac gag gtg atg cac cat gcc gtt ggc gca aac cga atg tgc	96
Tyr Pro His Glu Val Met His His Ala Val Gly Ala Asn Arg Met Cys	
20 25 30	
aag tgt gat gcc ccg gca ggg aac gcc gaa acc tcc gcc gac aga gag	144
Lys Cys Asp Ala Pro Ala Gly Asn Ala Glu Thr Ser Ala Asp Arg Glu	
35 40 45	
cag agt cac act ctc gat gag ttg acc cat cag ttg cac atg ctg cag	192
Gln Ser His Thr Leu Asp Glu Leu Thr His Gln Leu His Met Leu Gln	
50 55 60	
caa gcc tac gac acc ggc atg ggt cgt gtc gat gac gtg atg gag gac	240
Gln Ala Tyr Asp Thr Gly Met Gly Arg Val Asp Asp Val Met Glu Asp	
65 70 75 80	
atg gac gac ctg tcc cac agg atc gcc gac cac gag aag gaa cac tgt	288
Met Asp Asp Leu Ser His Arg Ile Ala Asp His Glu Lys Glu His Cys	
85 90 95	
aag aag tat aga gag ttc cag tgc ggt ggt gac cat cca aag tgc atc	336
Lys Lys Tyr Arg Glu Phe Gln Cys Gly Gly Asp His Pro Lys Cys Ile	
100 105 110	
tcg aac ctc ctc gtc tgc gac ggt gac aac gac tgt gac aat gga gct	384
Ser Asn Leu Leu Val Cys Asp Gly Asp Asn Asp Cys Asp Asn Gly Ala	
115 120 125	
gat gag gct cgt tgt gat gtg ctc acc gag gct ggt agt agt tgg act	432
Asp Glu Ala Arg Cys Asp Val Leu Thr Glu Ala Gly Ser Ser Trp Thr	
130 135 140	
ggt act gtg gtc tac gat cac tgc acc aag cgt cgc cca gag acc atg	480
Gly Thr Val Val Tyr Asp His Cys Thr Lys Arg Arg Pro Glu Thr Met	
145 150 155 160	
aag ctc agc atc aag agc gtg gat acc gta ccc ttc ttc acc acc cac	528
Lys Leu Ser Ile Lys Ser Val Asp Thr Val Pro Phe Phe Thr Thr His	
165 170 175	
ccc aag gtc cgc ggt acc gtg ctt atg gag aag cac acc aag gac tac	576
Pro Lys Val Arg Gly Thr Val Leu Met Glu Lys His Thr Lys Asp Tyr	
180 185 190	
agc gag gtc atc aac gag ccg gtc tct ggc tac tgg agc agc gcc gat	624
Ser Glu Val Ile Asn Glu Pro Val Ser Gly Tyr Trp Ser Ser Ala Asp	
195 200 205	
agg agc gcc gct atg ccc ccg gac agc gcc ggt cac ctt ggc ttt gtc	672
Arg Ser Ala Ala Met Pro Pro Asp Ser Ala Gly His Leu Gly Phe Val	
210 215 220	

tgc atc ttc cac ggc cac gac cac gac acc tgc act ggt ctc ctc acc 720
 Cys Ile Phe His Gly His Asp His Asp Thr Cys Thr Gly Leu Leu Thr
 225 230 235 240

aag ggc aag gtc aca gat gcc tgc gcc gag ttc acc ttc cac agg gat 768
 Lys Gly Lys Val Thr Asp Ala Cys Ala Glu Phe Thr Phe His Arg Asp
 245 250 255

taa 771

<210> 12
 <211> 256
 <212> PRT
 <213> Arenicola marina

<400> 12
 Met Lys Ser Tyr Val Leu Val Cys Cys Leu Val Val Gly Ala Val Ala
 1 5 10 15

Tyr Pro His Glu Val Met His His Ala Val Gly Ala Asn Arg Met Cys
 20 25 30

Lys Cys Asp Ala Pro Ala Gly Asn Ala Glu Thr Ser Ala Asp Arg Glu
 35 40 45

Gln Ser His Thr Leu Asp Glu Leu Thr His Gln Leu His Met Leu Gln
 50 55 60

Gln Ala Tyr Asp Thr Gly Met Gly Arg Val Asp Asp Val Met Glu Asp
 65 70 75 80

Met Asp Asp Leu Ser His Arg Ile Ala Asp His Glu Lys Glu His Cys
 85 90 95

Lys Lys Tyr Arg Glu Phe Gln Cys Gly Gly Asp His Pro Lys Cys Ile
 100 105 110

Ser Asn Leu Leu Val Cys Asp Gly Asp Asn Asp Cys Asp Asn Gly Ala
 115 120 125

Asp Glu Ala Arg Cys Asp Val Leu Thr Glu Ala Gly Ser Ser Trp Thr
 130 135 140

Gly Thr Val Val Tyr Asp His Cys Thr Lys Arg Arg Pro Glu Thr Met
 145 150 155 160

Lys Leu Ser Ile Lys Ser Val Asp Thr Val Pro Phe Phe Thr Thr His
 165 170 175

Pro Lys Val Arg Gly Thr Val Leu Met Glu Lys His Thr Lys Asp Tyr
 180 185 190

Ser Glu Val Ile Asn Glu Pro Val Ser Gly Tyr Trp Ser Ser Ala Asp
 195 200 205

Arg Ser Ala Ala Met Pro Pro Asp Ser Ala Gly His Leu Gly Phe Val
 210 215 220

Cys Ile Phe His Gly His Asp His Asp Thr Cys Thr Gly Leu Leu Thr
 225 230 235 240

Lys Gly Lys Val Thr Asp Ala Cys Ala Glu Phe Thr Phe His Arg Asp
245 250 255

```
<210> 13
<211> 376
<212> DNA
<213> Arenicola marina
```

```
<220>  
<221> CDS  
<222> (1) .. (375)  
<223>
```

[illegible]

```
<210> 14
<211> 125
<212> PRT
<213> Arenicola marina
```

```

<400> 14
Gly Pro Leu Gln Arg Leu Leu Val Lys Thr Gln Trp Asn Lys Val Tyr
1          5          10          15
Gly Thr Ser Lys Val Arg Asp Glu Ala Gly His Val Leu Trp Lys Ala
          20          25          30

```

Ile Phe Ala Gln Asp Pro Glu Thr Arg Ala Leu Phe Lys Arg Val Asn
 35 40 45
 Gly Asp Asp Ile Tyr Ser Pro Glu Phe Met Ala His Ser Ala Arg Val
 50 55 60
 Leu Gly Gly Leu Asp Ile Ala Ile Ser Leu Leu Asp Asn Gln Ala Asp
 65 70 75 80
 Leu Asp Val Ala Leu Ala His Leu His Val Gln His Val Glu Arg His
 85 90 95
 Ile Pro Thr Arg Tyr Phe Asp Leu Phe Lys Asn Ala Leu Met Glu Tyr
 100 105 110
 Ala Pro Ser Ala Leu Gly Arg Cys Phe Asp Lys Asp Ala
 115 120 125

<210> 15
 <211> 288
 <212> DNA
 <213> Arenicola marina

<220>
 <221> CDS
 <222> (1)..(288)
 <223>

<400> 15
 tgc gga ccc ctt cag cgc ctg aag gtc aag cgc cag tgg gct gag gct 48
 Cys Gly Pro Leu Gln Arg Leu Lys Val Lys Arg Gln Trp Ala Glu Ala
 1 5 10 15
 tat gga agc gga aac agc agg gag gaa ttc gga cac ttc atc tgg tcc 96
 Tyr Gly Ser Gly Asn Ser Arg Glu Glu Phe Gly His Phe Ile Trp Ser
 20 25 30
 cat gtc ttc cag cac tcg cct gct gcc cgc gac atg ttc aag cgc gtc 144
 His Val Phe Gln His Ser Pro Ala Ala Arg Asp Met Phe Lys Arg Val
 35 40 45
 cgc ggt gac aac atc cac acc cca gca ttc atg gcc cac gcc acc cgt 192
 Arg Gly Asp Asn Ile His Thr Pro Ala Phe Met Ala His Ala Thr Arg
 50 55 60
 gtg ctc ggt gga ctc gac atg tgc att gcc ctt ctc gat gat gaa ccc 240
 Val Leu Gly Gly Leu Asp Met Cys Ile Ala Leu Leu Asp Asp Glu Pro
 65 70 75 80
 gtt ctg aac acg cag ctc gct cat ctt gcc aag caa cac gaa acc cgt 288
 Val Leu Asn Thr Gln Leu Ala His Leu Ala Lys Gln His Glu Thr Arg
 85 90 95

<210> 16
 <211> 96
 <212> PRT
 <213> Arenicola marina

<400> 16

Cys Gly Pro Leu Gln Arg Leu Lys Val Lys Arg Gln Trp Ala Glu Ala
 1 5 10 15

Tyr Gly Ser Gly Asn Ser Arg Glu Glu Phe Gly His Phe Ile Trp Ser
 20 25 30

His Val Phe Gln His Ser Pro Ala Ala Arg Asp Met Phe Lys Arg Val
 35 40 45

Arg Gly Asp Asn Ile His Thr Pro Ala Phe Met Ala His Ala Thr Arg
 50 55 60

Val Leu Gly Gly Leu Asp Met Cys Ile Ala Leu Leu Asp Asp Glu Pro
 65 70 75 80

Val Leu Asn Thr Gln Leu Ala His Leu Ala Lys Gln His Glu Thr Arg
 85 90 95

<210> 17

<211> 360

<212> DNA

<213> Arenicola marina

<220>

<221> CDS

<222> (1)..(360)

<223>

<400> 17

aag gtg aag cac caa tgg gtg cag gtg tac agc ggc cat ggt tac gag 48
 Lys Val Lys His Gln Trp Val Gln Val Tyr Ser Gly His Gly Tyr Glu
 1 5 10 15

cgt gag gcg ttc ggc aga gag gtc ttc ctc gag atg tac aac cag gca 96
 Arg Glu Ala Phe Gly Arg Glu Val Phe Leu Glu Met Tyr Asn Gln Ala
 20 25 30

ccc aag gcc aag gac ctc ttc acc agg gtc agg ggc gag aac gtc ttc 144
 Pro Lys Ala Lys Asp Leu Phe Thr Arg Val Arg Gly Glu Asn Val Phe
 35 40 45

tcc ccc gag ttc gga gcc cac atg gtc cgt gtg ctc gga ggt ctc gac 192
 Ser Pro Glu Phe Gly Ala His Met Val Arg Val Leu Gly Gly Leu Asp
 50 55 60

atg tgc atc gct ctg ctg tcc gat gac acc gtc ctc aac gcc cag ctt 240
 Met Cys Ile Ala Leu Leu Ser Asp Asp Thr Val Leu Asn Ala Gln Leu
 65 70 75 80

gct cac ctc agc acg cag cac aag gac cgt gga atc ccc aac gag tac 288
 Ala His Leu Ser Thr Gln His Lys Asp Arg Gly Ile Pro Asn Glu Tyr
 85 90 95

ttc gat gtg gtg aag gtc gcc ctc atg aag gtc gtc ccc ggc cac gtt 336
 Phe Asp Val Val Lys Val Ala Leu Met Lys Val Val Pro Gly His Val
 100 105 110

tca cac ttc gat atc ggc gcg tgg
 Ser His Phe Asp Ile Gly Ala Trp
 115 120

360

<210> 18
 <211> 120
 <212> PRT
 <213> Arenicola marina

<400> 18
 Lys Val Lys His Gln Trp Val Gln Val Tyr Ser Gly His Gly Tyr Glu
 1 5 10 15
 Arg Glu Ala Phe Gly Arg Glu Val Phe Leu Glu Met Tyr Asn Gln Ala
 20 25 30
 Pro Lys Ala Lys Asp Leu Phe Thr Arg Val Arg Gly Glu Asn Val Phe
 35 40 45
 Ser Pro Glu Phe Gly Ala His Met Val Arg Val Leu Gly Gly Leu Asp
 50 55 60
 Met Cys Ile Ala Leu Leu Ser Asp Asp Thr Val Leu Asn Ala Gln Leu
 65 70 75 80
 Ala His Leu Ser Thr Gln His Lys Asp Arg Gly Ile Pro Asn Glu Tyr
 85 90 95
 Phe Asp Val Val Lys Val Ala Leu Met Lys Val Val Pro Gly His Val
 100 105 110
 Ser His Phe Asp Ile Gly Ala Trp
 115 120

<210> 19
 <211> 390
 <212> DNA
 <213> Arenicola marina

<220>
 <221> CDS
 <222> (1)..(390)
 <223>

<400> 19
 tgt tgc agt ata gag gac cgc aag gag gtc cag acg ctg tgg agt gag
 Cys Cys Ser Ile Glu Asp Arg Lys Glu Val Gln Thr Leu Trp Ser Glu
 1 5 10 15 48
 atc tgg agt gcc cag ttc act ggt cgc cgt gtc cag gtt gcc cag gct
 Ile Trp Ser Ala Gln Phe Thr Gly Arg Arg Val Gln Val Ala Gln Ala
 20 25 30 96
 gtg ttc gag gac ctc ttc cgc cgc gac ccc gag tcc aag aac ctg ttc
 Val Phe Glu Asp Leu Phe Arg Arg Asp Pro Glu Ser Lys Asn Leu Phe
 35 40 45 144

aag cgc gtc aat gtt gac gac atg aac agc ccc gaa ttc cac gct cac	192
Lys Arg Val Asn Val Asp Asp Met Asn Ser Pro Glu Phe His Ala His	
50 55 60	
tgc atc cgt gtt gtc aac ggt ctt gac acc gtg atc ggt ctc ctt gac	240
Cys Ile Arg Val Val Asn Gly Leu Asp Thr Val Ile Gly Leu Leu Asp	
65 70 75 80	
gac ccc gac acc ctg aag tcc cag ctc gag cac ttg gcc cag cag cac	288
Asp Pro Asp Thr Leu Lys Ser Gln Leu Glu His Leu Ala Gln Gln His	
85 90 95	
aag gag cgt gat ggc atc cac aag acc cac ttc gac gag atg tcc cac	336
Lys Glu Arg Asp Gly Ile His Lys Thr His Phe Asp Glu Met Ser His	
100 105 110	
gcc ttc ggc gcc gtc atg ccc cag gtc agc agc tgt ttc aac ccc gat	384
Ala Phe Gly Ala Val Met Pro Gln Val Ser Ser Cys Phe Asn Pro Asp	
115 120 125	
gca tga	390
Ala	

<210> 20
 <211> 129
 <212> PRT
 <213> Arenicola marina

<400> 20
 Cys Cys Ser Ile Glu Asp Arg Lys Glu Val Gln Thr Leu Trp Ser Glu
 1 5 10 15
 Ile Trp Ser Ala Gln Phe Thr Gly Arg Arg Val Gln Val Ala Gln Ala
 20 25 30
 Val Phe Glu Asp Leu Phe Arg Arg Asp Pro Glu Ser Lys Asn Leu Phe
 35 40 45
 Lys Arg Val Asn Val Asp Asp Met Asn Ser Pro Glu Phe His Ala His
 50 55 60
 Cys Ile Arg Val Val Asn Gly Leu Asp Thr Val Ile Gly Leu Leu Asp
 65 70 75 80
 Asp Pro Asp Thr Leu Lys Ser Gln Leu Glu His Leu Ala Gln Gln His
 85 90 95
 Lys Glu Arg Asp Gly Ile His Lys Thr His Phe Asp Glu Met Ser His
 100 105 110
 Ala Phe Gly Ala Val Met Pro Gln Val Ser Ser Cys Phe Asn Pro Asp
 115 120 125
 Ala

<210> 21
 <211> 20

<212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> A, G, C or T

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> A, G, C or T

<400> 21
 gartgyggnc cnttrcarcg

20

<210> 22
 <211> 21
 <212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<400> 22
 ctctctctcct ctctctcttcc t

21

<210> 23
 <211> 17
 <212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> A, G, C or T

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> A, G, C or T

<400> 23
 tgygggnathc tncarcg

17

<210> 24
 <211> 18
 <212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> inosine

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> A, G, C or T

<400> 24
 aargtnaarc anaactgg

18

<210> 25
 <211> 20
 <212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<400> 25
 tgytgyagya thgargaycg

20

<210> 26
 <211> 20
 <212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> A, G, C or T

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> A, G, C or T

<400> 26
 aargtnatht tyggnagrga

20

<210> 27
 <211> 20
 <212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> A, G, C or T

<220>
 <221> misc_feature
 <222> (18)..(18)
 <223> A, G, C or T

<400> 27
 garcaycart gyggnggnga

20

<210> 28
 <211> 21
 <212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<220>
 <221> misc_feature
 <222> (4)..(4)
 <223> A, G, C or T

<220>
 <221> misc_feature
 <222> (7)..(7)
 <223> A, G, C or T

<400> 28
 ccangcntcy ttrtcraagc a

21

<210> 29
 <211> 19
 <212> DNA
 <213> artificial sequence

<220>
 <223> PCR primer

<220>
 <221> misc_feature
 <222> (2)..(2)
 <223> A, G, C or T

<220>
 <221> misc_feature
 <222> (8)..(8)
 <223> A, G, C or T

<220>
 <221> misc_feature
 <222> (11)..(11)
 <223> A, G, C or T

<220>
 <221> misc_feature
 <222> (14)..(14)
 <223> A, G, C or T

<400> 29
antgygggncc nctncarcg

19

<210> 30
<211> 18
<212> DNA
<213> artificial sequence

<220>
<223> PCR primer

<220>
<221> misc_feature
<222> (4)..(4)
<223> A, G, C or T

<220>
<221> misc_feature
<222> (7)..(7)
<223> A, G, C or T

<400> 30
ccangcnccd atrtcraa

18

<210> 31
<211> 20
<212> DNA
<213> artificial sequence

<220>
<223> PCR primer

<220>
<221> misc_feature
<222> (3)..(3)
<223> A, G, C or T

<220>
<221> misc_feature
<222> (6)..(6)
<223> A, G, C or T

<400> 31
cangnycrc trtttraarca

20